

8/pets
1020 Rec'd PCT/PTO 14 JUN 2005Language pack for a communication terminal

The present invention relates to terminals comprising a
5 predictive editor application for entering text. Such a
terminal can be a cellular or cordless phone, a personal
digital assistant (PDA) or a communicator. The editor is
used for editing text for message handling, phonebook
editing and searching, etc. The invention further relates
10 to a method of compressing data in a mobile terminal
comprising the steps in which a language dependent
dictionary is stored.

BACKGROUND ART

15 EP 1031913 discloses a mobile terminal having processor
means controlling the display in accordance with the
operation of the keypad. A selectable predictive editor
program generates an output containing words matching a
20 received string of ambiguous key strokes. The predictive
editor program has a number of associated vocabularies
including at least one language dependent dictionary and
a dictionary receiving user defined inputs.

25 Language depended dictionaries such as stored in the
mobile terminal known from EP 1031913 needs to contain
sufficient words to be able to handle the most commonly
used words. Each language dependent dictionary (e.g. the
English one) requires 100 kbyte memory space. The
30 terminal usually stores several language dependent
dictionaries for a plurality of natural languages, e.g.
because the users are bi-lingual, or because the same
software is used in countries with different official
languages. The language dependent dictionaries and the
35 disambiguating software may therefore use up to 600
Kbyte of memory space.

Mobile terminals as known from EP1031913 have a number of language data sets, so-called language packs, stored thereon, that are used by the operating system (that is to say the menu control application) of the terminal to display text on display, e.g. in connection with menu control of the terminal. As a standard three to seven language packs for different natural languages are stored in the mobile terminal, so that users can choose which language they wish to use on the terminal, and so that the same software package can be used for mobile terminals sold in various countries in which different languages are spoken. A plurality of these language packs also represents considerable amount of data to be stored on the mobile terminal.

Data storage capacity is limited on mobile terminals due to the desire to construct these devices compact and light. Other applications on the mobile terminals also require data storage capacity, and therefore, it is desirable to use the available data storage capacity efficiently.

DISCLOSURE OF THE INVENTION

25

On this background, it is an object of the present invention to provide a mobile terminal of the kind referred to initially, that uses the available data storage capacity more efficiently. This object is achieved by a mobile terminal comprising a display; means for entering text; a predictive editor program for generating an output containing words completing a received string of unambiguous keystrokes or matching a received string of unambiguous key strokes, the predictive editor program has a number of associated vocabularies including at least one language dependent

dictionary; at least one further program that is capable of compressing text data by replacing words with references to the language dependent dictionary and/or capable of decompressing text data by retrieving words from the language dependent dictionary using references to the language dependent dictionary.

Most text data is stored using the 7 bit ASCII code that can represent 2^8 or 256 characters with each character coded by a unique 8 bit representation. Thus, a five character word placed in data storage requires 35 bits of information. In text transmission, a start bit, a stop bit are normally added to the character code for each character. Thus, that same five character word requires 5 times 10, i.e. 50 bits of information when it is transmitted.

With the present invention whole words are directly coded as references to the language dependent dictionary rather than of coding each of the characters that make up the particular alphanumeric word. This allows for reduction of storage space and decreases text transmission time. For example, for a vocabulary set containing N words, one can code the N words with a binary code having $\log_2 N$ bits. A choice of 14 bits per word results in a vocabulary set of 2^{14} or 16,384 words. Since each word in the vocabulary set consists of 14 bits regardless of the character count of the word, a significant reduction in storage space can be realized using this coding technique. A reduction will also be realized in text transmission time.

Preferably, the further program identifies words in data that is represented as a sequence of characters drawn from an alphabet in an input data block and processes it into an output data block in which character combinations that correspond to words contained in said language